Description of the MHS Health Level 7 Radiology for Public Health Surveillance

Technical Document NMCPHC-EDC-TD-5-2014

By Gosia Nowak
EpiData Center Department
March 2014

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**			

1. REPORT DATE *(DD-MM-YYYY)* 14-01-2014	2. REPORT TYPE Technical Document	3. DATES COVERED *(From - To)* July 2009 - December 2013

4. TITLE AND SUBTITLE Description of the MHS Health Level 7 Radiology for Public Health Surveillance	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Gosia Nowak	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Navy and Marine Corps Public Health Center EpiData Center Department 620 John Paul Jones Circle, Suite 1100 Portsmouth, VA 23708	8. PERFORMING ORGANIZATION REPORT NUMBER NMCPHC-EDC-TD-5-2014

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Navy and Marine Corps Public Health Center EpiData Center Department 620 John Paul Jones Circle, Suite 1100 Portsmouth, VA 23708	10. SPONSOR/MONITOR'S ACRONYM(S) NMCPHC
	11. SPONSOR/MONITOR'S REPORT NUMBER(S) NMCPHC-EDC-TD-5-2014

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release. Distribution is not limited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The EpiData Center Department (EDC) at the Navy and Marine Corps Public Health Center (NMCPHC) evaluated the Health Level 7 (HL7) data source for its usefulness in health surveillance activities. This technical document provides a history of the HL7 radiology database and its contents, explains the creation of radiology records, describes the pathway of data from healthcare provider to the EDC, provides a detailed descriptions of all variables within the database, and assesses the database's strengths and limitations. Given an understanding of the strengths and limitations of the data, HL7 radiology data have proven to be a valuable source of health information for surveillance purposes. The data allow the creation of a timeline of events corresponding to a specific disease occurrence. Furthermore, data are received in a timely fashion, allowing for near-real-time surveillance of diseases.

15. SUBJECT TERMS

Health Level 7 (HL7), Radiology, Surveillance

16. SECURITY CLASSIFICATION OF: U			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Gosia Nowak
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U	UU	28	19b. TELEPHONE NUMBER *(include area code)* 757-953-0979

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

NAVY AND MARINE CORPS PUBLIC HEALTH CENTER
PREVENTION AND PROTECTION START HERE

Abstract

The EpiData Center Department (EDC) at the Navy and Marine Corps Public Health Center (NMCPHC) evaluated the Health Level 7 (HL7) data source for its usefulness in health surveillance activities. This technical document provides a history of the HL7 radiology database and its contents, explains the creation of radiology records, describes the pathway of data from healthcare provider to the EDC, provides a detailed descriptions of all variables within the database, and assesses the database's strengths and limitations. Given an understanding of the strengths and limitations of the data, HL7 radiology data have proven to be a valuable source of health information for surveillance purposes. The data allow the creation of a timeline of events corresponding to a specific disease occurrence. Furthermore, data are received in a timely fashion, allowing for near-real-time surveillance of diseases.

Table of Contents

Executive Summary

Project Background

The EpiData Center (EDC) at the Navy and Marine Corps Public Health Center (NMCPHC) was funded to evaluate the Health Level 7 (HL7) Radiology data source for its usefulness in public health surveillance. This technical document is the result of those efforts. The radiology dataset contains records documenting radiographic exams since 6 July 2009.

Public Health Surveillance Applications

Radiology data add a unique layer to the EDC's surveillance efforts. Because these data are not limited to physician diagnoses or laboratory confirmed cases, they can provide supporting information for laboratory confirmed; physician diagnosed, or presumptively treated cases. The greatest value of radiology data for the Navy and Marine Corps currently lies in disease-specific procedures. However, many symptoms and treatments are not specific to a particular disease or condition. Radiologic procedures depend on the suspect disease but may be general in the type of procedure performed (e.g. ultrasound, chest x-ray, bone x-ray). The procedure type does not indicate specific disease but results may support clinical diagnosis or treatment. The use of radiological procedures may be dependent on provider practice, severity of symptoms, medical or family history. Data on radiology transactions, therefore, can improve the robustness of surveillance systems based on lab results and/or ICD-9 coded records.

Key Fields for Public Health Surveillance

Specific key fields for Public Health Surveillance are included in the data: SPONSOR ID, family member prefix (FMP), SERVICE, REQUESTING FACILITY, and PERFORMING FACILITY.

True duplicates are defined as records in which all fields are exactly the same. After true duplicates are eliminated, the data can be analyzed by unique patient, test, or record. *Unique patients* are identified in the HL7 radiology data through a combination of SPONSOR ID and FMP, this combination creates a unique identifier that can be used to track individual patients through all radiology records. A *unique order* is defined as all records associated with each radiographic exam. A *unique record* is defined as each record associated with each radiographic exam for each patient.

Strengths

Several of the data fields of interest are complete but the completeness of the database as a whole continues to be assessed. The analysis showed that data were collected in the Composite Health Care System (CHCS) from the majority of the DOD military treatment facilities (MTFs). The timeliness of data processing is within the acceptable range for the Navy disease surveillance activities, typically two days.

Limitations

It is currently not clear whether Defense Health Services Systems (DHSS) captures all CHCS radiology transactions. Further work is needed to compare HL7 radiology records to other data sources in order to estimate completeness. The radiology data only include MTFs that have CHCS servers, and exams performed at forward deployed clinics, contracted managed care support clinics, and other MTFs that do not use CHCS are not captured. Incomplete demographic information (e.g., unspecified marital status, race, or ethnicity) can limit the generalizability of these data to specific minority groups. Extra precautions need to be taken when extrapolating data to larger populations and when comparing disease rates and trends among the military to non-military populations.

Project Background

The EpiData Center (EDC) at the Navy and Marine Corps Public Health Center (NMCPHC) was funded to evaluate the Health Level 7 (HL7) radiology data source for its usefulness in public health surveillance. This technical document is a result of those efforts. The radiology dataset contains records documenting radiographic exams performed at a military treatment facility (MTF). Records for all Department of Defense (DOD) military service members (Army, Navy, Marine Corps, Air Force, Coast Guard, and US Public Health Service), overseas civilian personnel, Tri-Care eligible dependents, and others who receive their radiological exams at a military MTF are included in this dataset. The following document describes the original observations on the data fields, some basic frequencies, the cleaning rules implemented for usability, and other comments relevant to the use of these data for surveillance.

Initial evaluation of the dataset involved one sample extract received by the EDC from the Defense Health Services System (DHSS). This was a very small dataset used to analyze the structure, completeness, and distribution of the entire data set. The extract was received for Message Dates between 1 September 2013 and 31 December 2013. Descriptive analysis on these data included frequency distribution of demographic fields, evaluation of null or invalid values for key fields used in surveillance, and understanding data structure in the extracts received compared to the structure as data is entered into the Composite Health Care System (CHCS). The extract was reviewed and analyzed in order to determine applicability for surveillance and modify the data structure to more accurately address the disease surveillance needs of the Navy and Marine Corps, as well as other services. The current data archive dates back to 6 July 2009.

Data Origination and Flow Process

The radiology dataset includes all exams that are performed at an MTF. There are several mechanisms of entry that can occur. The most common process followed is described below along with notable exceptions.

A radiology order is initially entered into the CHCS system by the ordering physician. The radiology department receives the order via CHCS and verifies it. If clarification is needed, staff may contact ordering physician for more information. When the radiologist or radiology technician completes the radiological procedure, he/she enters the procedure information (e.g. image type, result text), completes the record and saves it in the local CHCS system. If results are edited upon verification, edits are made in the CHCS record. The radiologist has the ability to cancel orders with physician approval. Each time a record is canceled, changed, edited, or reordered, a new record in CHCS is generated.

The HL7 radiology data are limited to radiologic exams at an MTF facility that uses CHCS. If orders are entered into CHCS and not completed (exam is not performed), records are not present in the HL7 radiology dataset. Exams completed in a purchased care facility are also not included in this dataset. Forward deployed clinics, battalion aid stations, and shipboard clinics do not participate in the CHCS system, and records are not received from these locations.

Public Health Surveillance Applications

Radiology data add a unique layer to the EDC's surveillance efforts. Because these data are not limited to physician diagnoses or laboratory confirmed cases, they can provide supporting information for laboratory confirmed; physician diagnosed, or presumptively treated cases. The greatest value of radiology data for the Navy and Marine Corps currently lies in disease-specific procedures. However, many symptoms and treatments are not specific to a particular disease or condition. Radiologic procedures depend on the suspect disease but may be general in the type of procedure performed (e.g. ultrasound, chest x-ray, bone x-ray). The procedure type does not indicate specific disease but results may support clinical diagnosis or treatment. The use of radiological procedures may be dependent of provider practice, severity of symptoms, medical or family history. Data on radiology transactions, therefore, can improve the robustness of surveillance systems based on lab results and/or ICD-9 coded records.

Current surveillance methods in the EDC include monitoring HL7 lab results, ICD-9 codes in clinical encounter records, and outpatient/inpatient pharmacy transactions. Consequently, surveillance methods are largely disease-specific, but this specificity depends on ICD-9 coding practices in local MTFs, timeliness of lab testing, ability to accurately flag lab tests of interest, and disease specific treatment regimens. The use of radiology data will greatly improve the surveillance of certain infectious diseases, such as tuberculosis, because other data on these diseases are so limited by delayed lab test results and gross inaccuracies in ICD-9 coding.

Potential use of HL7 radiology records is not limited to surveillance. Data on radiological procedures can fill critical gaps in the military's ability validate specific diagnoses such as active tuberculosis infection, pneumonia, hyperemia, emphysema and aspiration. Coupled with laboratory and encounter data, disease management guidelines can be evaluated. Finally, these data may provide valuable insight into clinical practice and atypical disease presentation.

Data Structure and Analysis

HL7 radiology data are retrieved by the EDC in a standard, pipe-delimited flat file from DHSS. Each column within the data file is a fixed variable and each row should contain a unique record. Each person can have more than one distinct record, if they have multiple radiology procedures or updates to their exams. Each exam ordered is associated with a unique record (row). The variable fields are designed to ease analysis, except for the free text fields, which require the use of wildcards or search terms.

Key Fields for Public Health Surveillance

Defining Duplicates

Within the HL7 radiology dataset there are several ways in which duplicate records can be identified. Duplicate rules described here should be checked against project objectives to ensure the applicability. *True duplicates* are defined as records in which all fields are exactly the same. Records meeting this criterion should be eliminated. There are three types of unique records described here, which are most relevant to public health surveillance purposes: unique record, unique person, and unique order.

Unique Record

Each record that remains after removing true duplicates is considered a unique record. There is at least one variable value different than all other records in the database.

Unique Person

Patients are identified in the HL7 radiology data through a combination of SPONSOR ID and FMP. This combination creates a unique identifier that can be used to track individual patients through all radiology records. There is a variable that includes the PATIENT ID, however, the EDC has previously determined that the field is not complete, consistent, or reliable as a source of identifying patients within or across databases. It is important to note that it is possible for individuals to have two separate SPONSOR IDs over time. For example, if the child of a sponsor becomes active duty, then that child will have his/her own SPONSOR ID. Each unique patient can have multiple exam orders in the HL7 radiology data.

Unique Order

A unique order is defined as all records associated with a single specific radiology exam. Each exam ordered is assigned an ORDER NUMBER. ORDER NUMBERs may be reused, however it is unlikely that a person would receive the same order number more than once. The combination of SPONSOR ID, FMP, and ORDER NUMBER can be used to identify unique orders within the radiology dataset. Each unique order can have multiple records within the HL7 radiology data.

Corrected Records

The EDC currently receives records that are completed or a record that is designated with a RESULT STATUS of "F" (final). If a record is corrected (result status of "C" = amended), an additional record is generated. The information contained in the original record is included in the updated record. Additional/corrected information is appended to the SET ID RESULT TEXT field (original findings remain in this field, as well), and when the message is present, the message date/time and DHSS LOAD DATE time are updated by CHCS. If a record indicated a change is present then that record should be considered in the analysis instead of the initial record. In 0.06% of orders the record is corrected more than once. When this occurs a new record is generated in the same manner as before, adding the additional information to the SET

ID RESULT TEXT field of the first corrected record. If an order has more than one corrected order the record with the most recent MESSAGE DATE and MESSAGE TIME should be considered in analysis.

Strengths

Timeliness

DHSS includes several date fields in the dataset provided to the EDC: ARRIVAL DATE, DHSS LOAD DATE, MESSAGE DATE, OBSERVATION DATE, ORDER DATE, ORDER EFFECTIVE DATE, RESULT REPORT DATE, and TRANSACTION DATE. A timeline of dates is provided in Appendix A. To assess the timeliness of the data, the RESULT REPORT DATE (date the result was certified) was compared to the MESSAGE DATE (date the HL7 message was generated by CHCS) to estimate the time between the exam completion and the receipt of data at DHSS. The MESSAGE DATE was also compared to the DHSS LOAD DATE in order to determine the time between HL7 message generation at the local CHCS host and DHSS data parsing of the HL7 message into the database design.

For almost all records (99%), an HL7 message was generated the same day as the result report date. After generation, it took about two days for the message to be processed by DHSS (87%). Based on processing of the data at DHSS, NMCPHC receives these data within approximately two days, though this time estimate needs be verified. The delay in data receipt indicates that the timeliness of reporting is within acceptable ranges for the Navy disease surveillance activities. Future analysis and assessment should define lag times in relation to particular MTFs or disease outcomes of interest.

Completeness

Records are received from the majority of shore-based fixed MTFs connected to the CHCS, but gaps in data may exist. Gaps in data may occur due to server failure at location or due to functional errors. It is believed that these data represent at least 90% of all filled completed radiology procedures in CHCS.   The completeness of individual fields varies and the characteristics of each are described in detail in the field observations section that appears later in this document. In general, some fields of particular interest, such as SPONSOR ID, FMP, and SERVICE are highly populated due to the business rules of CHCS.

Limitations

Completeness
The HL7 infrastructure at DHSS was built using pilot funds but as of March 2009 is a functional program. Initially, a temporary network was created to capture HL7 messages when they were sent from the CHCS host to the Denver feed node. Up until the program became formal, no back-up system existed. When the feed node fails, HL7 messages may be lost and those that have been sent may not be retrievable unless the network outage was planned for in advance. Gaps may exist in the data received at NMCPHC, though are limited. Several of the identified data fields of public health interest are highly populated, but others are not. The completeness of each data field, as described later in this document, should be considered before its use in analysis.

Inclusion
The data only includes MTFs that utilize the CHCS servers. Forward deployed clinics, contracted managed care support clinics, and other MTFs that do not use CHCS are not captured in these data unless the order is filled by a radiology department that uses CHCS. The CHCS is not used to order or process radiological exams onboard ships.

Generalizability
Incomplete demographic information (e.g. unspecified MARITAL STATUS, RACE, or ETHNICITY) can limit the generalizability of these data to specific minority groups. Demographic information not provided in this database can be supplemented with other available personnel databases.

Comparability
These data are generated from the radiographic exam records of a highly specific patient population – military service members and other military beneficiaries – which differs from the general U.S. population in many ways, including average age, gender distribution, physical fitness, and health status. Further, this population has universal access to medical care, which is not true of the U.S. population. These differences limit the comparability to the general US population. Extra precautions need to be taken when extrapolating data to larger populations and also when comparing the disease rates and trends of the military and non-military populations.

All Data Fields (Variables)

The following section describes frequency distributions run on all fields within the HL7 radiology database, based on snippet of records between 01 September 2013 and 31 December 2013. Any problems that arose in relation to data values were addressed with DHSS and resolved to the best possible conclusion. The data fields of most interest include SPONSOR ID, FMP, SERVICE, REQUESTING FACILITY, PERFORMING FACILITY, and other fields that are necessary for the EDC's planned surveillance activities.

Automatically Populated Fields

There are several types of automatically populated fields in the radiology data.

When a facility registers within the CHCS, several variables are created, which identify the facility: PERFORMING DMIS ID, PERFORMING FACILITY, PERFORMING FACILITY SERVICE, PERFORMING WORK CENTER, RADIOLOGY SITE, REQUESTING DMIS ID, REQUESTING FACILITY, REQUESTING FACILITY SERVICE, and REQUESTING WORK CENTER.

When DHSS compiles the data from the CHCS server, two fields are automatically populated: DHSS LOAD DATE and DHSS LOAD TIME.

Each patient or beneficiary is registered in the Defense Eligibility Enrollment Reporting System (DEERS) under the SPONSOR ID, which feeds into the CHCS system. When a patient presents at a medical facility, the SPONSOR ID (usually the Social Security Number) is entered and their name is chosen from a drop down list. The following patient demographic fields are automatically populated after this selection, if they were entered when the patient was registered in DEERS: DATE OF BIRTH, ETHNICITY, FMP, Gender, Marital Status, Patient Category, Patient ID, Race, Service, and SPONSOR ID. If these data are not present in the system, a designated unknown value is entered, and therefore there are no missing values in these fields. Registration is completed and records updated when the sponsor reports to a new command and selects an MTF. Administrative personnel at the MTF have the ability to edit records at the time of visit.

As records are created, edited, and completed, the date and time variables are created by the CHCS system. These variables can be changed, if necessary, by the user, but this change is not common practice.

MSG DATE, MSG TIME, and MSG SENDING FACILITY are created and assigned when the message (record) is sent to the CHCS server.

Formatting

Several variables in the radiology data contain numerical values. A few of these fields may contain leading zeros that would affect analysis if lost: SPONSOR ID, PATIENT ID, FMP,

PERFORMING FACILITY DMIS ID, and REQUESTING FACILITY DMIS ID. In order to maintain the data integrity these fields should be imported in character format.

Generation of Facility Information

When each facility registers with CHCS, the facility name is created. Each record generated from the location will have the same facility name. If the facility name was entered incorrectly (e.g. if it was misspelled), it will be consistently incorrect in all records from that facility. Within each facility there are a variety of work centers that can generate radiology records. The work center variable is a free text field that the ordering physician fills during order generation.

The EDC has provided DHSS with an official DOD Defense Medical Information System Identifier (DMIS ID) list. This list is used to create a DMIS ID for each record based on the information contained in the facility name field. Once records have been assigned a DMIS ID, additional fields describing the facility are created: DMIS FACILITY NAME and FACILITY SERVICE. If the DMIS ID is missing, either because the facility name was missing or a correct match was not made, these variables are also missing. Furthermore, a secondary quality assurance check is performed on the raw data once it is received at NMCPHC. Records with null values in the DMIS ID field are identified. For those records, an algorithm based on the REQUESTING and/or PERFORMING FACILITY NAME fills in the DMIS ID.

The DMIS ID is listed for both the requesting and the performing facility. REQUESTING FACILITY DMIS ID indicates which facility placed the order for the exam. PERFORMING FACILITY DMIS ID indicates the facility at which the exam was performed. Information about the performing or requesting facility may be useful for certain projects.

Field Observations (in alphabetical order):

Appendix A, which contains a timeline and a brief description of the date variables in the dataset, should be consulted to better understand the relationship of the dates included.

ARRIVAL DATE

This field contains the date of patient arrival. This field is formatted: YYYYMMDD. Values for this variable are missing in 0.6% of records.

ARRIVAL TIME

This field contains the time of patient arrival. This field is formatted: HHMM, and values range from 0000-2359. Values for this variable are missing in 0.6% of records.

CLINICAL COMMENTS

The CLINICAL COMMENTS field contains relevant clinical information noted during the exam. This is a free text field and not required for each order. There is no additional information provided within this field (values are missing) in 83% of records.

CPT CODE DATA

The CPT CODE DATA field contains CPT-4 code data for radiology procedures.

It has four components separated by "\": CPT Code, Modifier, Status, and Date/Time of CPT action. It is structured to facilitate analysis but could be separated if necessary. Values for this variable are missing in less than 1% of records.

DATE OF BIRTH
The DATE OF BIRTH field (DOB) is formatted YYYYMMDD. Birthdates are valid in over 99% of records (invalid records are those with default values including '19000101' and '19010101', along with those indicating a day or month value of '00').

DHSS LOAD DATE
DHSS LOAD DATE indicates the date when DHSS loads the data from the central CHCS server. The field is used to determine the timeliness of reporting and to identify lags in reporting times from certain MTFs. The format is YYYYMMDD and there are no records with missing values.

DHSS LOAD TIME
Time component of the DHSS LOAD DATE field, and is formatted: HHMM. The values present in the data are 0300, 1000, 1600, and 2000, and there are no records with missing values.

ETHNICITY
ETHNICITY is a numeric field with six possible values: 1=Hispanic, 2=South East Asian, 3= Filipino, 4=Other Asian Pacific Islander, 9=Other, and Z=Unknown.

There are no missing values in this field, though 93% of all records indicate a value for Other or Unknown. These results indicate that the ETHNICITY field is not consistently completed. This lack of specificity potentially limits the ability to identify disease trends and burden in minority groups unless the data are supplemented from other personnel information.

EXAM NUMBER
The EXAM NUMBER is an eight digit numeric value that identifies each order and its associated observations. There are no missing values for this field.

FMP
FMP is the family member prefix, which designates the relationship of the patient to the sponsor. The distribution of FMP was as expected, with the highest frequencies indicating children (01-19), sponsors (20), and spouses (30-39). There are no missing values for this field. A full list of FMP codes is available for analysis.

The FMP field is automatically populated in the CHCS system. When the SPONSOR ID is entered, the possible patient names are listed. When the appropriate patient name is selected, the FMP field is completed.

GENDER
There are three values possible for GENDER: M=Male, F=Female, and X=Unknown. There are no missing values for GENDER and less than 1% of records have a value indicating unknown.

IMAGING TYPE

The IMAGING TYPE field contains information on the imaging modality. It is formatted as ##\TEXT. It may be possible to use the initial numbers (i.e. ##) to limit records to a specific type of imaging modality, though definition for these values are not currently available. Less than 0.1% has missing values for this field.

INTERPRETING RADIOLOGIST

This field includes the name of the interpreting radiologist. It is formatted as "Last Name, First Name, Middle Initial (where available)". Less than 1% has missing values for this field.

MARITAL STATUS

There are nine values for MARITAL STATUS: A=Annulled, D=Divorced, I=Interlocutory Decree, L=Legally Separated, M=Married, N=Never Married, S=Single/Not Married, W=Widow or Widower, and Z=Unknown.

There are no missing values for MARITAL STATUS. However, 42% of all records are categorized as unknown. Single/Not Married and Married account for the largest percentages of those with known values.

MEPRS CODE

The MEPRS CODE is a four letter code that indicates where within the MTF the person received treatment. The first letter indicates the most general area and translates as: A=inpatient, B=outpatient, C=dental, D=ancillary, E=support, F=special programs, and G=readiness. It is advised that an up-to-date list of all possible codes be obtained. The field is useful for tracking where people are seen within the MTF. For instance, it can indicate ambulatory care, special dialysis clinics, the maternity ward, and so on, which can affect the interpretation of the data.

The majority of records have a MEPRS CODE that begins with B (89%). There are no records with missing values for this variable.

MSG DATE

This field is automatically assigned in CHCS when the order is completed in the system and sent to the CHCS server. The date approximates the TRANSACTION DATE but it can vary between locations. Some MTFs send messages in batches, therefore the date portions may not correlate to the actual TRANSACTION DATE. This field is formatted: YYYYMMDD. There are no missing values for this variable.

MSG ID

MESSAGE ID is an alphanumeric code assigned to each batch of messages based on when the message is sent from CHCS to the server. The MESSAGE ID is not unique to each record; each batch of messages is assigned one MESSAGE ID. The format of MESSAGE ID varies by MTF and includes numbers, letters and/or a numeric code that identifies the MTF or it can identify the function of the message (e.g. RESCHED-057342). There are no records with missing values for this variable.

MSG SENDING FACILITY

This field identifies the facility that sends the message to DHSS through the CHCS system. This field allows analysts to identify and track problems that arise in the transfer of messages from the MTFs through DHSS to the EDC. There are no records with missing values for this variable.

MSG TIME

This field is automatically assigned in CHCS when the order is completed in the system and sent to the CHCS server. The time approximates the TRANSACTION TIME but it can vary based on location. Some MTFs send messages in batches, therefore the time portions may not correlate to the actual TRANSACTION TIME. This field is formatted: HHMM. There are no missing values for this variable.

NO OF CPT CODES

This field indicates the NUMBER OF CPT CODES received in a given message. The range of values present in current data is 1 to 12, and the value for this field is missing in 2% of records.

OBSERVATION DATE

This field contains the date of the radiology exam. This field is formatted: YYYYMMDD. The value for this field is missing in less than 1% of records.

OBSERVATION TIME

This field contains the time of the radiology exam. This field is formatted: HHMM, and values range from 0000-2359. The value for this field is missing in less than 1% of records.

ORDER DATE

The ORDER DATE field contains the date of the radiology procedure. This field is formatted YYYYMMDD and does not include missing values.

ORDER EFFECTIVE DATE

The ORDER EFFECTIVE DATE is the date that the order takes effect or is supposed to take effect. The ORDER EFFECTIVE DATE is formatted YYYYMMDD and does not include missing values.

ORDER EFFECTIVE TIME

The ORDER EFFECTIVE TIME is the time that the order takes effect or is supposed to take effect. The ORDER EFFECTIVE TIME is formatted as HHMM, and ranges from 0000-2359. There are no records with missing values for this variable.

ORDER NOTES/COMMENTS

The ORDER NOTES/COMMENTS is a free text field that contains information describing the diagnosis or reason for the procedure. This field may consist of multiple segments concatenated to include the complete text of the order notes or comments. Concatenated information will be separated by the '~' symbol.

ORDER NUMBER

ORDER NUMBER is a numeric code of eleven digits (xxxxxx-xxxxx) unique to each order but not unique for each record. An order can have multiple records that correspond to changes made to the order (e.g., changes in result interpretations). All changes appear as individual records with the same ORDER NUMBER. It is a plausible way to track a patient but it is not useful for identifying unique records. There are no missing values in the ORDER NUMBER field.

ORDER STATUS

ORDER STATUS field designates the status of an order. Values include: CA=cancelled, CM = Completed, DC=discontinued, ER = error, HD=on hold, IP=in process unspecified, RP = replaced, and SC = scheduled. This field indicates the status of the order when the message was created. All records present indicate a completed record. The value for this field is missing in less than 1% of records.

ORDER TIME

The ORDER TIME field contains the time of the radiology procedure. This field is formatted HHMM, and values range from 0000-2359. This field does not include missing values.

ORDERING PROVIDER

The ORDERING PROVIDER field indicates the name of the physician ordering the radiology exam. It has four components separated by "\": Last Name, First Name, Middle Name (or initial), Suffix. It is structured to facilitate analysis but could be separated if necessary. There are no records with missing values for this variable.

PATIENT CATEGORY

The PATIENT CATEGORY CODE (PatCat) is an alphanumeric code that indicates the patient's relationship to the uniformed services. The first letter of the code refers to the branch of service of the Sponsor (A=Army, B=National Oceanic and Atmospheric Administration, C=Coast Guard, F=Air Force, K=other beneficiaries of the federal government, M=Marine Corps, N=Navy, P=US Public Health Service, R=NATO Recipient). It is followed by two digits corresponding to the status of the Sponsor, as well as the patient's relationship. For example: A11=Army Active Duty Member, A41=Army Dependant of Active Duty, etc. A complete list is available from the DOD for analysis. Values for this field are missing in less than 1% of records.

PATIENT ID

The PATIENT ID is intended to serve as a unique identifier for each patient. In the data dictionary provided by DHSS, it states that the PATIENT ID is the patient's SSN, when available. There are no records with missing values in this field.

Based on the EDC evaluation, the PATIENT ID should not be used to identify patients. This field may contain an ID number assigned by the facility to a patient instead of the patient's SSN, and therefore cannot be used to match with other data sources or follow a patient across multiple

facilities. It is recommended that SPONSOR ID and FMP be used together to create a unique identifier instead.

PERFORMING DMIS FACILITY NAME

This field is the text translation of the DMIS ID provided in the PERFORMING DMIS ID field. The value for this variable is missing in 16% of records.

PERFORMING DMIS ID

The PERFORMING DMIS ID is a four digit code that identifies the performing facility. This code allows for grouping of MTFs based on geographic location, as well the ability to identify parent/child relationships between installations. The value for this variable is missing in 1% of records.

PERFORMING LOCATION FACILITY

The PERFORMING LOCATION FACILITY field in CHCS indicates the name of the MTF where the exam is completed. There are no records where the value of this variable is missing.

PERFORMING FACILITY SERVICE

The PERFORMING FACILITY SERVICE field indicates the service branch that the MTF that filled the prescription. The possible values for this field are: A=Army, C=Coast Guard, F=Air Force, and N=Navy. There are missing values for this variable in 16% of records.

PERFORMING LOCATION WORK CENTER

The PERFORMING LOCATION WORK CENTER field indicates the specific work location that completes the radiology exam. There are no records with missing values for this variable.

PROCEDURE NAME

PROCEDURE NAME is a free text field that contains a description of the radiology procedure. This field does not contain information about procedure results. There are no missing values for this field.

RACE

There are six possible values for RACE; C=White, M=Asian or Pacific Islander, N=Black, R=American Indian or Alaskan Native, X=Other, and Z=Unknown. There are no missing values for RACE. However, approximately 52% of the records are classified as Other or Unknown.

RAD EXAM STATUS

This field indicates the status of the requested radiology exam. This field is formatted: Unique ID/Exam Status. The majority of records have the value '2\COMPLETE' in this field, other values present account for less than 1% of all records. Other values present include: '8\AMENDED', 16\EXAMON, and '17\REPORT'. A full list of values is not present in available documentation. There are missing values in less than 1% of records.

RECORD TYPE

DHSS enters a value of "RAD" for all radiology records. The value does not change between extracts and there are no missing records. The field has an analytic value if different HL7 datasets are used in a single analysis as it will serve to identify the source of each individual record.

REQUESTING DIVISION NAME

The REQUESTING DIVISION NAME field is derived from the second portion of the requesting location description and is the name of the "work center" that requested the radiology exam. This field is free text and contains no missing values.

REQUESTING DMIS FACILITY NAME

This field is the text translation of the DMIS ID provided in the REQUESTING DMIS ID field. There are missing values for this variable in 18% of records.

REQUESTING DMIS ID

The REQUESTING DMIS ID is a four digit code that identifies the facility that ordered the prescription. This code allows grouping of MTFs based on geographic location, as well as to identify parent/child relationships between installations. There are missing values for this variable in 3% of records.

REQUESTING FACILITY NAME

The REQUESTING FACILITY NAME field is derived from the first portion of the requesting location description and is the name of the facility where the order originated. This field is free text and contains no missing values.

REQUESTING FACILITY SERVICE

The REQUESTING FACILITY SERVICE field indicates the service branch of the facility that requested the prescription. Possible values for this field are: A=Army, C=Coast Guard, F=Air Force, and N=Navy. There are no records indicating Coast Guard. There are missing values for this variable in 18% of records.

RESULT STATUS

The RESULT STATUS field contains the result status of the radiology observation. Possible values include: O=unacknowledged, I=pending, S=scheduled, P=intermediate, C=amended, R=uncertified, F=completed and certified, and X=order canceled. The majority of records present (>98%) indicate a complete and certified record; the remaining records are amended or uncertified. The value for this field is missing in less than 1% of records.

RESULTS REPORT DATE

The RESULTS REPORT DATE is the date that the report is approved. The RESULTS REPORT DATE is formatted YYYYMMDD. The value for this field is missing in less than 1% of records.

RESULTS REPORT STATUS

The RESULTS REPORT STATUS field includes the status of the radiology report. Possible values include: A = amended, PD = problem draft, and V = verified. More than 99% of the records had a value of V (verified). There are missing values in less than 1% of records.

RESULTS REPORT TIME

The RESULTS REPORT TIME is the time that the report is approved. The RESULTS REPORT TIME is formatted as HHMM, and ranges from 0000-2359. The value for this field is missing in less than 1% of records.

SERVICE

The SERVICE field refers to the service branch of the Sponsor. The value is determined from the first component of the PATIENT CATEGORY field and the values are the same. Therefore, there are less than 1% of records with missing values for SERVICE, the same records missing PATCAT codes.

SET ID RESULT TEXT

The SET ID RESULT TEXT field is a concatenation of the SET ID and the RESULT TEXT FIELDS.

The SET ID field identifies the logical order of arrival of each segment of the radiology report. Each radiology report can span several segments of the HL7 record. Each RESULT TEXT segment has been concatenated with the SET ID field to provide the order the text was written. All segments were then ordered and concatenated into one field to provide the complete radiology report.

The RESULT TEXT field includes the text of the radiology report. Each radiology report can span several segments of the HL7 record. Each segment has been concatenated with the SET ID field to provide the order the text was written. All segments were then ordered and concatenated into one field to provide the complete radiology report.

SPONSOR ID

The SPONSOR ID field corresponds to the Social Security Number (SSN) of the sponsor and is in the format of #########, with no dashes. There are no missing values for this field. SPONSOR ID is not sufficient to serve as a unique identifier for each patient but it can be used in conjunction with the FMP to identify a unique patient.

SUPERVISING RADIOLOGIST

This field includes the name of the supervising radiologist. It is formatted as "Last Name, First Name, Middle Initial (where available)". All values for this field are missing.

TRANSACTION DATE

The TRANSACTION DATE is the date that the order enters the CHCS system. The TRANSACTION DATE is formatted YYYYMMDD. The value for this field is missing in less than 1% of records.
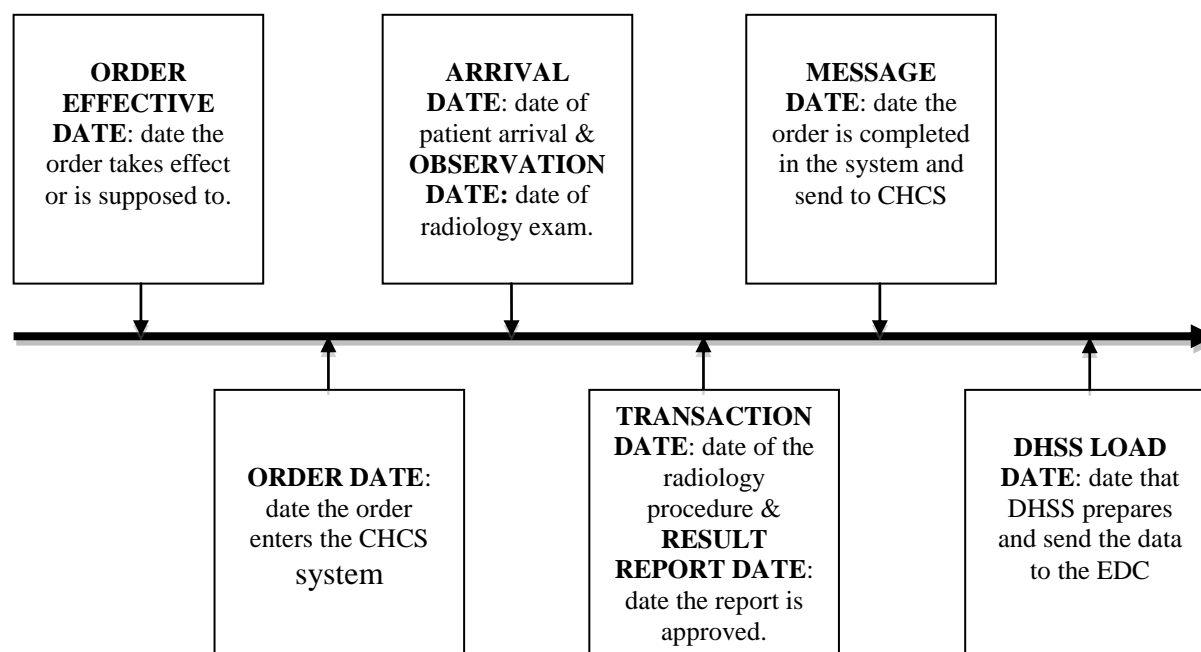
TRANSACTION TIME

The TRANSACTION TIME field represents the time component of the TRANSACTION DATE formatted as HHMM, and ranges from 0000-2359. The value for this field is missing in less than 1% of records

Appendix A: Dates

There are eight date variables in the HL7 radiology data received by the EDC. Analyses were completed based on existing records to determine the relationship between the date variables.

All records have the same date for two sets of variables. ARRIVAL DATE has the same value in all records as OBSERVATION DATE. TRANSACTION DATE has the same value in all records as RESULT REPORT DATE.

ORDER EFFECTIVE DATE: date the order takes effect or is supposed to.	**ARRIVAL DATE**: date of patient arrival & **OBSERVATION DATE:** date of radiology exam.	**MESSAGE DATE**: date the order is completed in the system and send to CHCS

ORDER DATE: date the order enters the CHCS system	**TRANSACTION DATE**: date of the radiology procedure & **RESULT REPORT DATE**: date the report is approved.	**DHSS LOAD DATE**: date that DHSS prepares and send the data to the EDC

The order shown in the above timeline is the typical order, and some records may have records in a different order. Many of the dates have the same value as other fields. The table below shows the percent of records that are the same between each variable and the next. DHSS prepares and sends data (DHSS LOAD DATE) to the EDC one day after the order is completed and sent to the CHCS system (MESSAGE DATE) in 97% of records.

Date 1	Date 2	Percent Same Value
ORDER EFFECTIVE DATE	ORDER DATE	83%
ORDER DATE	ARRIVAL DATE/ OBSERVATION DATE	84%
ARRIVAL DATE/ OBSERVATION DATE	TRANSACTION DATE/ RESULT REPORT DATE	68%
TRANSACTION DATE/ RESULT REPORT DATE	MESSAGE DATE	99%

POINT OF CONTACT

Navy and Marine Corps Public Health Center
XXX Division
EpiData Center Department
WWW.NMCPHC.MED.NAVY.MIL/

Contact Name
757.953.####
contact.name@med.navy.mil